Adventures with Jukeboxes

Mike Martin and Kalyani Rengarajan
Data Distribution Laboratory

Introduction

The Data Distribution Laboratory (DDL) maintains a collection of scientific CD-ROM disks from all government agencies. Most titles are obtained through reciprocal distribution agreements (we provide copies of titles we produce, in return USGS, National Earthquake Information Center, EROS Data Center and others provide us their new titles). Others we beg, borrow or steal and occasionally purchase (the National Geophysical Data Center is particularly stingy). The collection has now grown to well over 200 disks, representing about 150 gigabytes of scientific data. The DDL publishes an illustrated catalog of Government Scientific CD-ROM titles, which can be obtained by contacting the PDS Operator on Internet at PDS_OPERATOR@JPLPDS.JPL.NASA.GOV.

Although this immense collection can be stored in several inches of shelf space in CD binders, it is not as accessible as we would like. We have been searching for several years for a CD jukebox that would house the collection and provide Internet access to data files on the disks. One device, developed by Kubik Technologies, has been demonstrated at trade shows for many years, but has always been just a few months away from real availability as a product. In October, 1992 we were finally able to persuade Kubik Enterprises, Inc. to provide us with an evaluation unit.

Jukebox Description

The Kubik CDR-240M Compact Disk Changer (Figure 1) holds 240 CD-ROMs and can be equipped with from one to four CD-ROM readers. The Toshiba XM-3301 player has been installed in most units produced to date, though Kubik advertises that other readers can be installed. Future deliveries will probably utilize the Toshiba XM-3401 double-speed (300 kilobytes per second) reader. The jukebox is currently priced at $20,000.

The first jukebox was delivered in mid-October, and a Kubik representative flew to JPL to perform the installation. Unfortunately, when the equipment was unpacked it was found that all four CD players were damaged. The Kubik mechanism requires the use of a heavy steel CD caddy and the caddies were not properly secured before shipment. Another jukebox was delivered in mid-November, with caddies secured and no damage occurred.

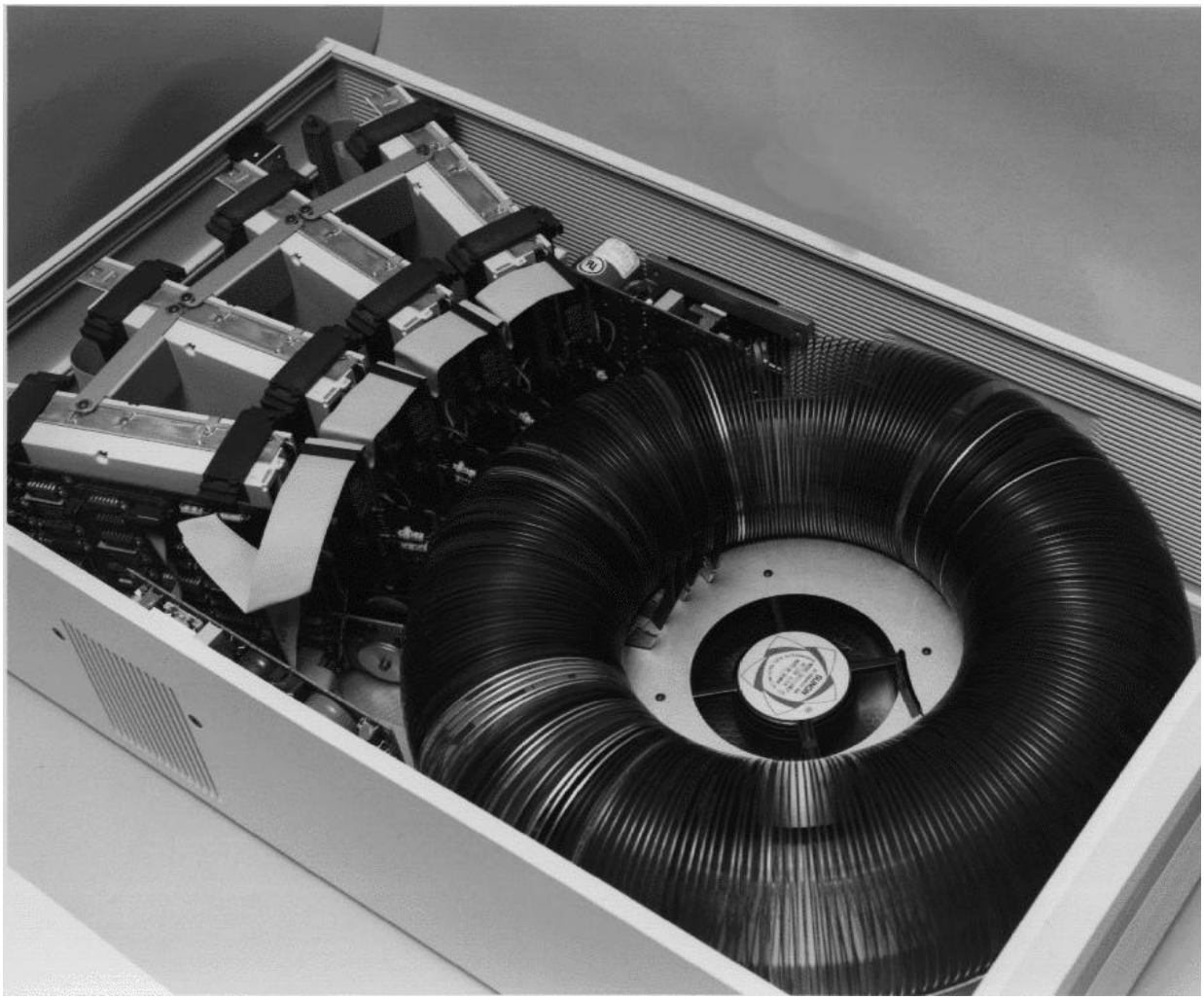

Figure 1. Kubik Jukebox with scientific CD-ROM collection

The jukebox was supplied with IBM-PC software running under Windows 3.1. The software presents a scrolling menu of the titles of CD's in the jukebox (which are manually entered when disks are loaded in the jukebox). A DOS command can also be associated with each disk, so that an application can be automatically started when the disk is mounted. This is a very useful capability in a reference library environment, but not as important with our scientific data collections. The software was developed using Toolbook, a hypercard-like application development system from Asymetrics. The 'driver' software for UNIX workstations which was mentioned in Kubik literature did not exist, which meant that we were essentially starting from scratch in our effort to utilize the jukebox on a Sun workstation.

Hardware Interface

The Kubik has two interfaces, a serial port that controls the jukebox mechanism and a SCSI port which provides access to the CD readers. Our initial efforts to install the jukebox were hampered by interface problems. It took more than a week to

successfully communicate over the serial interface, mainly due to our unfamiliarity with Sun serial port protocols and our inability to find anyone with expertise in this area. Once we were able to talk to the machine we found that the controls are very simple, consisting of characters codes representing actions to be performed. For example, "P1S221M" tells the jukebox to mount (M) the disk in slot 221 (S221) in drive 1 (P1). Once the disk is loaded in the reader, a SCSI mount command must be issued to identify the volume to the Sun file system. This initially required superuser privileges, but we have since received a special program which allows unprivileged device mounts and dismounts (available via anonymous ftp at cdrom.com in /pub/mount.c). The CD reader drivers we are using to access the drives are from Young Minds. To dismount the disk two steps are required. First a SCSI eject command must be sent to the reader, then the command "P1U" is sent to the serial port. The jukebox remembers (in non-volatile CMOS memory) what slot the disk came from and will return it to the proper position. The command "S221L" command instructs the jukebox to load/unload disks (one at a time) through a port in the front of the machine. The jukebox can be bulk loaded by removing the cover and inserting disks, but in normal operation the cover must be on the device, since some of the controls are light sensitive. Other commands include a sequential load which puts away the current disk in a specified drive then mounts the disk in the next slot number; a status query "?" which indicates what disks are in what drives; and the "CW" command to clear the non-volatile CMOS memory

The Sun workstation used to host the Kubik has two SCSI busses. The second bus was purchased with a large hard disk drive and we assumed could accommodate other devices. That was an incorrect assumption, apparently SCSI busses can be device specific and the vendor informed us that their bus did not support CD devices. Thus we were forced to completely reconfigure our workstation, moving all hard disk and tape drives to the second SCSI bus to free up the Sun SCSI bus for use with the Kubik. This involved about two weeks of effort. We were also informed by Sun technical support that their bus will reliably support only two CD-ROM readers, apparently because of the slow transfer rate. We ignored this advice and connected all four anyway and they seem to work, however it is possible that this architecture is responsible for some sporadic system problems we have encountered. It took at least a week to get a working version of the SCSI eject command, which was not included in the standard capabilities of the Young Minds CD drivers we are using and had to be extracted from a test program they provided.

Software Applications

Two simple applications have been developed for testing the jukebox system. First a 'torture test' program randomly selects disks and drives and mounts and unmounts them for a specified number of repetitions. We want to make sure the device is not prone to recurring failures which would require frequent maintenance.

Second, a retrieval application has been developed for selecting and mounting disks. The database which is used to produce the Catalog of Scientific CD-ROMs was exported

from Mac FileMaker Pro to an ASCII text file which was loaded into Sybase running on a UNIX workstation. A simple query menu was developed in Sybase to allow selection of disks by volume id, mission, target or title. Once a disk is selected by the user the Sybase program sends a request to our simple jukebox driver software to mount the specified volume on a drive. The user is then informed that the disk is available on the file system.

Test Results

Initial testing has resulted in an 'out of sync' condition on every reader in the jukebox. This condition requires physically resetting the drive, and is our greatest concern at this time. There also seem to be some idiosyncrasies in communicating with the serial interface, it does not always seem to respond per specification. The device should transmit a single acknowledgment character for each command received, but sometimes transmits extra acknowledgments. Also, when attempting to mount a disk in a drive that is busy the device should return an error code, but instead indicates that the command has been completed. Contrarily, sometimes a error code is returned even though the disk and drive are available. There also seem to be potential timing sensitivities, that is if commands are sent too quickly the controller hangs. There is a general problem that the commands do not time out, the failure to load a disk will result in infinite retries, essentially hanging the unit until it is manually reset.

Summary.

As of mid-March the evaluation unit has been returned to Kubik to be retrofitted with all new internal mechanisms to bring it up to the state of jukebox units currently being manufactured. Our plan is to run the torture test for the equivalent of 3 months of average use which we assume to be about 3,000 mounts. This would be a reasonable period between maintenance. If the unit holds up under this testing we will purchase it, if not we will work with Kubik to resolve the reliability problems. It seems clear that the CD reader caddy handling mechanisms are the most important threat to proper operation at this time. The mechanical positioning of the turntable has been cited as a potential point of failure, but we have seen no indication of this in our testing.

If we purchase the device, we then face the much bigger problem of determining the right user interface to a massive scientific collection of 150 gigabytes. There are many issues including:

- Should access be on a volume level or file level?
- How are drives allocated to users?
- How much time does each user get?
- Is there one global interface to the device or many different user interface applications dealing with a server?

As another consideration, many of the disks contain software which needs to be run to access the data. Ideally one would like to export these volumes to the user for Network File System (NFS) access, so users can access them as local devices attached to their own host computers.

We are currently investigating all these issues. We plan to explore the virtual volume approach, as is used in many optical jukebox systems (the AMASS software for example), where all the disks appear to be on-line in a giant file system. We also hope to apply the STELAR interface approach (see Information Systems Newsletter, Jan 1993), using WAIS as a front end to portions of the data base. It should be relatively easy to load a descriptive data base for the Planetary Data System (PDS) CD-ROM collection, since each disk contains a standard set of documentation with detailed data set and instrument descriptions. It would also be useful to tie the jukebox into the PDS Detailed Level Catalogs which reside at Geosciences and Imaging nodes, so that searches of the massive Viking, Voyager and Magellan image collections could result in immediate image access. We would appreciate any collaborative proposals from other tasks which are investigating these issues.

There are several other jukebox options on the horizon. A German firm is developing a 100-disk jukebox with a single reader. Young Minds is currently evaluating this jukebox. It is expected to be available this summer at a price on the order of $8,000. It is possible that this unit could be equipped with a CD-recorder and 100 blank disks, providing a 68 gigabyte storage system for less than $10K!. Of course the disks must be recorded a volume at a time, not file by file, so this would not be a competitor to magneto-optic jukeboxes at this time. 3rd Point Systems in Santa Monica is negotiating with magneto-optic jukebox vendors to develop a 2,000 disk jukebox system with 10 Pioneer DRM-604X readers included. The DRM-604X uses 6-disk cartridges and provides a quad-speed (600 kilobytes per second) read rate. Finally, Borett Automation Technologies, 818-597-8664 offers a versatile robotics device which can be used with many different types of storage devices and media. As another alternative, Peter Yee at Ames Research Center is installing 14 Pioneer Changers containing 84 CDs. The volumes can be accessed via Internet anonymous ftp at 'explorer.arc.nasa.gov'. For updates on the Kubik testing contact Kalyanir@michelle.jpl.nasa.gov.

Acknowledgments.

We would like to thank Stacey Teramae, George Karas, Dave Hecox and Steve Hughes at JPL for their hardware and software support. We thank Lyle Kerr and Jim Duncan of Kubik Technologies and Laslo Sipos of Kubik Enterprises, Inc. for providing technical support and for providing the evaluation unit.

Kubik Technologies, #110-13120 Vanier Place, Richmond, BC, 604-273-0400.

Kubik Enterprises, Inc., 18873 Allendale Avenue, Saratoga, CA 95070, 408-867-7969.

Young Minds, Incorporated, 1919 Orange Tree Lane, Suite 300, Redlands, Ca 92374, 713-335-1350.